DOCUMENT RESUME

ED 080 552 TM 003 057

AUTHOR Bratfisch, Oswald; And Others
TITLE Perceived Item-Difficulty in Three Tests of Intellectual Performance Capacity.
INSTITUTION Stockholm Univ. (Sweden). Inst. of Applied Psychology.
SPONS AGENCY Swedish Council for Social Science Research, Stockholm.
PUB DATE 72
NOTE 16p.; Reports from the Institute of Applied Psychology, University of Sweden, No. 29

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Complexity Level; *Intelligence Tests; *Item Analysis; *Perception; *Rating Scales; Statistical Analysis; Technical Reports
IDENTIFIERS Psychophysics

ABSTRACT

Three tests of intellectual performance capacity referring to factors V, S, and R, according to Thurstone's system of primary mental abilities were administered to a total number of 34 subjects. Immediately after finishing an individual item, subjects were asked to estimate the perceived difficulty of that item. The ratings were to be given on a symmetrical scale with 9 categories with verbal expression labels. A high correlation between the rank order of items according to estimated difficulty and the real item sequence was obtained in all three tests used. A linear relationship was found between estimated difficulty and standard scores corresponding to solution frequencies. A close correspondence was noticed between the widths and the levels of the ranges of the estimates on the one hand and the corresponding widths and levels of the standard score ranges on the other hand. Subjects who could solve an item correctly tended to estimate the difficulty of that item as lower than subjects who could not. (Author)

FILMED FROM BEST AVAILABLE COPY

REPORTS FROM THE INSTITUTE OF APPLIED PSYCHOLOGY
THE UNIVERSITY OF STOCKHOLM

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED 080552

TM 003 057

OSWALD BRATFISCH
GUNNAR BORG
STANISLAV DORNIČ

PERCEIVED ITEM-DIFFICULTY IN THREE TESTS OF INTELLECTUAL PERFORMANCE CAPACITY

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
G. Borg
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

No. 29, 1972

© 1972 Institute of Applied Psychology,
University of Stockholm,
Råsundavägen 101, 171 37 Solna, Sweden
Director and Editor: Gunnar Borg

PERCEIVED ITEM-DIFFICULTY IN THREE TESTS OF INTELLECTUAL PERFORMANCE CAPACITY*

Bratfisch, O., Borg, G., and Dornič, S. Perceived item-difficulty in three tests of intellectual performance capacity. Reports from the Institute of Applied Psychology, the University of Stockholm, 1972, No. 29. - Three tests of intellectual performance capacity referring to factors V, S, and R according to Thurstone's system of primary mental abilities were administered to a total number of 34 subjects. Immediately after finishing an individual item subjects were asked to estimate the perceived difficulty of that item. The ratings were to be given on a symmetric scale with 9 categories with verbal expression labels. A high correlation between the rank order of items according to estimated difficulty and the real item sequence was obtained in all three tests used ($r \geqslant 0.92$). A linear relationship was found between estimated difficulty and standard scores corresponding to solution frequencies. A close correspondence was noticed between the widths and the levels of the ranges of the estimates on the one hand and the corresponding widths and levels of the standard score ranges on the other hand. Subjects who could solve an item correctly tended to estimate the difficulty of that item as lower than subjects who could not.

Introduction

A series of investigations at our institute has been concerned with "subjective" difficulty of various human activities, i.e. with difficulty as perceived by the performing person himself. One of our starting points was the reflection that perceived difficulty of any given activity rather than corresponding "objective" measurements would be decisive for a person's feelings, attitudes, motivation, etc., concerning that activity.

Researchers do not appear earlier to have payed much attention to this problem area. Guilford and Cotzin report a study on perceived (felt) difficulty of judgement tasks in 1941, but relatively

---

* This investigation was supported by research grants from the Swedish Council for Social Science Research (Project numbers 439/71P and 350/71P). The authors are indebted to Dr. L. Hallsten for valuable comments on this manuscript.

little work seems to have been done in this field in the following two decades. Interest in the measurement of perceived difficulty was, however, taken by Borg (1961), whose study was followed by a series of investigations on the topic covering physical performance (e. g. Borg, 1962), motor skill (Bratfisch, Dornič & Borg, 1970), immediate memory (Borg, Bratfisch & Dornič, 1971a), visual search tasks (Borg, Bratfisch & Dornič, 1971b), and a test of reasoning ability (Bratfisch, Dornič & Borg, 1972). Theoretical, methodological and applied problems in connection with perceived difficulty have been discussed as well (Borg, Bratfisch & Dornič, 1971c).

The present study falls within the above outlined research program and concerns the perceived difficulty of three different kinds of intellectual tasks. Our main interest was in the questions (a) how is perceived difficulty related to measurements based on performance and (b) are there any differences between the estimates of difficulty of the three kinds of intellectual tasks.

The experiments

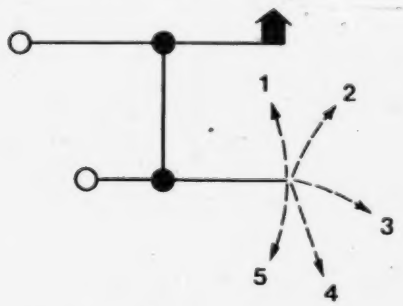
Items in three factor tests from a standardized intelligence battery (Delta Battery, 1969) regularly applied in connection with vocational guidance were used as stimuli in the present experiments. The tests refer to factors reasoning ability (R), spatial ability (S), and verbal comprehension (V) in accordance with the system introduced by Thurstone and Thurstone (1941) as indicated below:

(1) "Number series" (R)

1 2 4 8 16 32 - -

Find the rule by which the numbers are arranged and fill in the two numbers that are missing.

(2) "Levers" (S)



Indicate in which direction the end of the lever will move if you move the upper lever bar in the direction of the arrow.

(3) "Synonyms" (V)

BOY girl man lad guy
(translated from Swedish)

Underline one of the four words in small letters that means about the same as the word in capitals to the left.

The selection of tests was guided by the endeavour to cover intelligence factors which have been shown to account for a large part of general ability (e. g. Härnqvist, 1960). Thus also a bigger variety of intellectual activities was covered than had been done in earlier investigations on perceived difficulty, where mainly tests of reasoning ability were applied (e. g. Borg & Forsling, 1964; Borg, 1969).

Methods and experimental conditions

The three tests were administered to the subjects under standard conditions. During the performance part of the test the subjects were asked to estimate the perceived difficulty of the individual items immediately after they had finished an item. No information feedback about the correctness of a solution was given. Difficulty was to be estimated even when an item remained unsolved, i. e. even when a subject after having tried to solve an item "gave up". The estimates were to be given on a nine-grade symmetrical category scale. The categories were assigned verbal labels as follows: 1 - very, very easy; 2 - very easy; 3 - easy; 4 - rather easy; 5 - neither easy nor difficult; 6 - rather difficult; 7 - difficult; 8 - very difficult; 9 - very, very difficult. The scale was presented prior to the testing session and the subjects were carefully instructed to base the estimates on their immediate experience of difficulty regardless of the item sequence, as item sequence was not necessarily arranged according to "objective" difficulty. The experiments were carried out either individually or in small groups of 3 to 4 subjects and in one session.

Means and standard deviations of the experimental estimates together with z-values corresponding to the solution frequencies from a group of 100 vocational guidance clients of our institute are shown in Tables 1, 2, and 3, for the tests "Number series", "levers", and "Synonyms" respectively.

Table 1 Means (M) and standard deviations (s) of the experimental estimates in the test "Number series" as well as standard scores (z) corresponding to solution frequencies obtained from another group of 100 persons.

Item	M	s	z	Item	M	s	z
1	1.85	0.86	-1.405	13	6.19	1.44	+0.306
2	1.48	0.58	-2.326	14	5.15	1.43	+0.332
3	2.63	1.11	-1.341	15	5.96	1.53	-0.050
4	3.44	1.25	-0.643	16	7.11	1.16	+1.175
5	3.44	1.09	-0.332	17	6.67	1.00	+0.954
6	4.48	1.50	-0.706	18	6.26	1.58	+0.306
7	5.33	1.49	+0.176	19	7.63	1.08	+1.555
8	4.30	1.07	-0.878	20	7.11	1.28	+1.282
9	4.93	1.30	-0.842	21	7.33	1.42	+2.326
10	5.00	1.24	-0.306	22	7.70	1.20	+2.326
11	5.56	1.34	-0.253	23	7.67	1.33	+1.341
12	4.67	1.33	-0.675	24	8.15	0.86	+2.054
				25	8.78	0.51	+2.326

Table 2 Means (M) and standard deviations (s) of the experimental estimates in the test "Levers" as well as standard scores (z) corresponding to solution frequences obtained from another group of 100 persons.

Item	M	s	z	Item	M	s	z
1	4.24	1.00	-0.915	13	5.38	1.36	-0.279
2	4.33	1.06	-0.878	14	5.90	1.37	-0.126
3	4.86	1.39	-0.279	15	5.86	1.74	+0.279
4	4.95	1.53	-0.279	16	5.62	1.88	-0.202
5	4.76	1.14	-0.075	17	5.76	1.58	-0.279
6	5.05	1.20	-0.050	18	6.29	1.23	-0.100
7	4.76	0.94	-0.739	19	6.19	1.21	+0.126
8	4.86	1.42	-0.739	20	5.86	1.42	-0.050
9	5.38	1.24	-0.468	21	6.81	1.26	+1.126
10	5.38	1.43	-0.100	22	6.90	1.58	+0.332
11	6.14	1.01	+0.332	23	7.14	1.53	+0.202
12	5.33	1.15	-0.202	24	7.29	1.59	+0.279

Table 3 Means (M) and standard deviations (s) of the experimental estimates in the test "Synonyms" as well as standard scores (z) corresponding to solution frequencies obtained from another group of 100 persons.

Item	M	s	z	Item	M	s	z
1	1.33	0.56	-2.326	16	4.33	1.69	+0.176
2	1.91	1.10	-1.751	17	3.04	1.46	-1.476
3	1.58	0.50	-1.751	18	3.92	1.74	-0.995
4	2.83	1.34	-0.995	19	4.00	1.29	-0.995
5	2.79	1.06	-2.326	20	5.00	1.74	-0.583
6	3.04	1.34	-1.645	21	5.17	1.55	+0.228
7	3.29	1.20	-1.405	22	5.17	1.49	+0.151
8	4.08	1.41	-1.881	23	4.75	1.36	+0.643
9	3.62	1.41	-0.613	24	5.88	1.54	+0.413
10	3.79	1.86	-0.806	25	6.17	1.55	+1.341
11	3.75	1.82	+0.075	26	5.83	1.95	+1.341
12	4.46	1.32	-0.842	27	6.04	1.63	+0.413
13	3.92	1.64	-1.751	28	5.75	1.80	+0.279
14	3.87	1.54	-0.496	29	6.71	2.12	+0.100
15	4.83	1.79	-0.440	30	7.13	1.65	+1.341

Subjects

Altogether 34 subjects with high school education participated in the experiments. Twenty-seven of them participated in the experiment with the test "Number series", 21 in the one with the test "Levers", and 24 in the one with the test "Synonyms". Eight subjects were involved in all the experiments, 22 in two, and 4 in only one. The proportion of males to females was in all three experimental groups about 1:1. The age of the subjects ranged from 20 to 37 years with a median age of 23.

Results and discussion

Inter-individual variation in estimated difficulty

A number of investigations have shown that the inter-individual variation in estimates is largely depending on the scaling method applied (see e. g. Ekman & Künnapas, 1969); Bratfisch, Dornič & Borg, 1972). An analysis of the relation between means and the corresponding standard deviations can thus be regarded as an aspect of the reliability of the estimates when no other information (e. g. repeated estimated of the stimuli) is available.

In Figure 1A the standard deviations of the experimental estimates of each item in the test "Number series" are plotted against the corresponding means.* Figures 1B and 1C show the same kind of data for the tests "Levers" and "Synonyms", respectively.

An inverse U-shaped trend is roughly descriptive for the relation between standard deviations and mean estimates in Figure 1A, while the standard deviations are growing linearly with increasing means in Figure 1C. The data in Figure 1B show a considerable scatter, but a linear function can be said also here to describe the trend approximately from a parsimonious point of view. An inverse U-shaped relation between standard deviations and means is to be expected when the frequency distributions of the extreme stimuli on a scale with defined upper and lower boundaries are truncated by not permitting the subjects to place any stimuli below or above the boundaries. This was the case with "Number series" (Fig. 1A), but no such end effects were noticed with regard to "Levers" (Fig. 1B) and only in single cases with "Synonyms" (Fig. 1C). On the contrary, the means of the estimates concerning "Levers" cover categories 4 to 7 only, and those of "Synonyms" mainly categories 1 to 7, and no marked skewness of the individual item-distributions was noticed in either of these two tests with the exception of the first two stimuli of the test "Synonyms". With this kind of data the relation between standard deviations and means is expected to be by and large linear. Referring to the reasoning above the reliability of the data obtained in the three experiments can be regarded as satisfactory.

---

* When analyzing the frequency distributions of the items of the three tests they did not appear to be skewed or truncated (with but a few exceptions concerning "Number series" and "Synonyms") and it was decided thus to present means and standard deviations, though medians and semi-interquartile ranges were computed as well. No considerable difference, however, could be noticed when using either of these statistics.

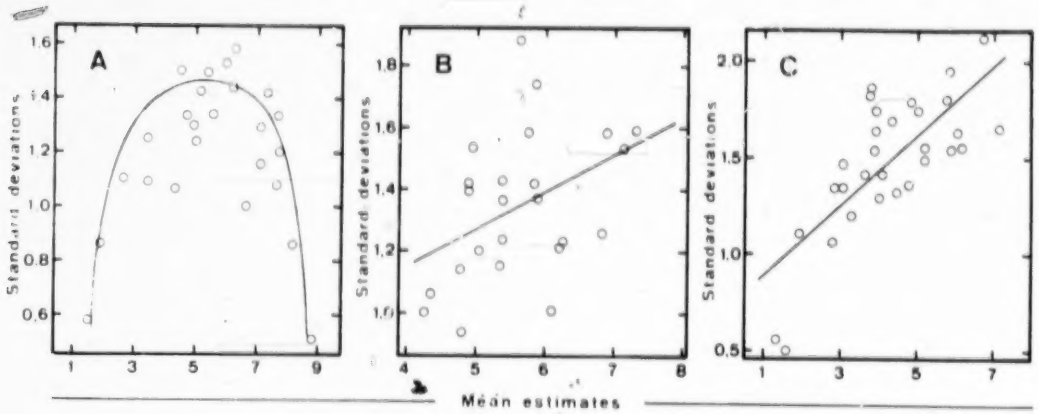


Fig. 1 Standard deviations plotted against arithmetic means of estimates. Diagram A shows data from "Number series", Diagram B those from "Levers", and Diagram C those from "Synonyms". The regression lines in Diagrams B and C were fitted mathematically, the curve drawn in Diagram A was fitted by eye.

Estimated and "objective" difficulty

Figure 2 shows means of the experimental estimates of difficulty plotted against the order of items in the tests. Diagram A represents the data from "Number series", Diagram B those from "Levers", and Diagram C those from "Synonyms". The close relationship between the sets of data is in all three graphs quite evident and is numerically confirmed by Spearman coefficients of rank-order correlation of 0.97, 0.92 and 0.92 for the data in Diagrams A, B, and C, respectively.

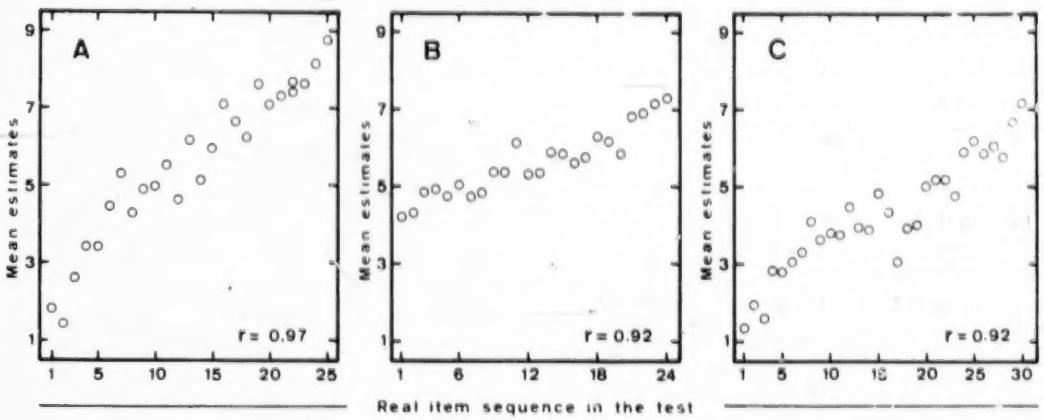


Fig. 2 Means of estimates plotted against the real order of items in the tests. Diagram A shows the data from "Number series", Diagram B those from "Levers", and Diagram C those from "Synonyms".

It could be argued that these close relationships are due to the fact that items were to be estimated in the order of occurrence in the test. However, similar investigations (Borg & Forsling, 1964; Borg, 1969; Bratfisch et al., 1972), where items were presented randomly and not according to item sequence in the test, showed rank-order correlation coefficients of about 0.90 between real item sequence in the test and rank order according to estimates of difficulty. That the coefficients of correlation are only slightly higher in the present investigation probably means that the order in which items are presented has but negligible effects on the estimates.

Figure 3 shows means of perceived difficulty as a function of standard scores (z-values), corresponding to the solution frequencies (p) of a group of 100 vocational guidance clients with the same level of education as the experimental groups (cf Tables 1, 2, and 3). The category scale applied is thus treated as a scale with equal intervals. The usage of z-values corresponding to solution frequencies from another group of 100 persons instead of those of the three experimental groups in Figure 3 is due to the fact that the standard error of a p-index also depends on the size of the sample population, which is rather small in the experimental groups. Diagram A in Figure 3 represents again the data from "Number series", Diagram B those from "Levers" and Diagram C those from "Synonyms".

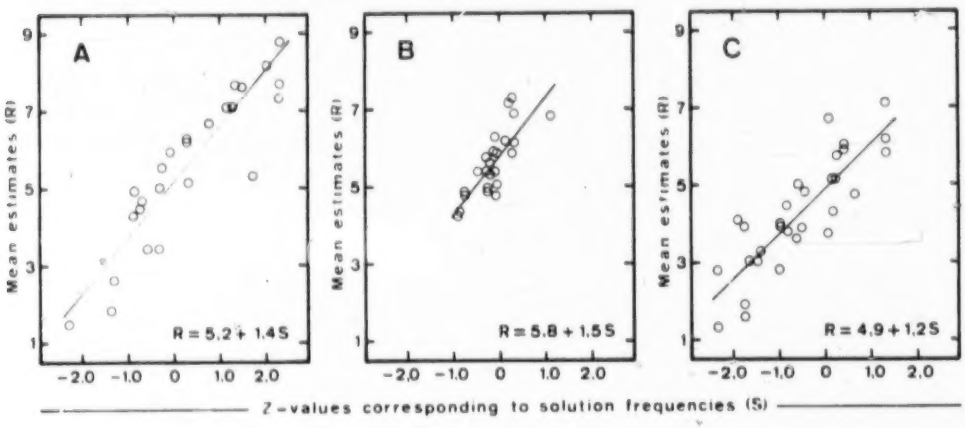


Fig. 3 Means of estimates related to standard scores (z-values). Diagram A shows data from "Number series", Diagram B those from "Levers", and Diagram C those from "Synonyms".

The slopes of the fitted regression lines are 1.4, 1.5, and 1.2, the Pearson coefficients of correlation are 0.94, 0.80, and 0.82, and the Spearman coefficients of rank order correlation between the rank order based on the mean estimates and the rank order according to the z-values are 0.95, 0.80, and 0.84 for the data in Diagrams A, B, and C, respectively. Rank-order coefficients of correlation of 0.90 between averaged estimates of difficulty and z-values as well as a by and large linear relationship has been found in similar investigations (Borg & Forsling, 1964, 1965; Bratfisch et al., 1972).

Apart from the close relationship between the sets of data, all the diagrams in Figure 3 show a close correspondence between the widths and the levels of the ranges of the estimates on the one hand and the

corresponding width and levels of the z-ranges on the other hand. An "objectively" medium difficult item is also estimated as medium difficult even if it is presented first when estimating, as the results from e. g. "Levers" show. A rather difficult item according to solution frequencies is also estimated as rather difficult (and not e. g. as very, very difficult) even if it is presented as one of the last items to be judged, as the results from both "Levers" and "Synonyms" show.

The striking differences in estimates between the three tests are hardly due to differences between the experimental groups, as the results from the following analysis show. Means of the experimental estimates of those 8 subjects participating in all three experiments were computed. These values are plotted against z-values in Figure 4. Diagram A shows again data from "Number series", Diagram B those from "Levers", and Diagram C those from "Synonyms".

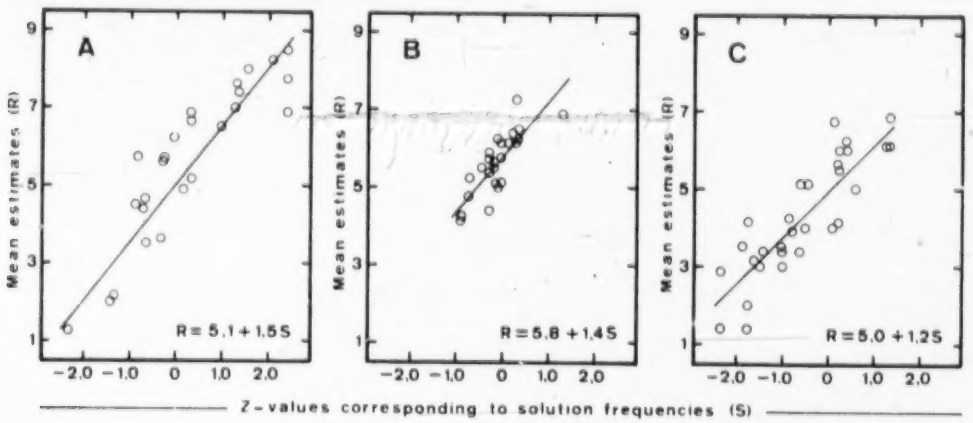


Fig. 4 Means of estimates of 8 subjects participating in all three experiments as related to standard scores (z-values). Diagram A shows data from "Number series", Diagram B those from "Levers", and Diagram C those from "Synonyms".

The results shown in Figure 4 are almost identical with those shown in Figure 3. In Figure 4 the slopes of the fitted regression lines are 1. 5, 1. 4, and 1. 2, the Pearson coefficients of correlation are 0. 95, 0. 81, and 0. 80, and the Spearman coefficents of rank order correlation between the rank order based on the mean estimates and the rank order according to the z-values are 0. 91, 0. 82, and 0. 87 for the data in Diagrams A, B, and C, respectively. The very close correspondence between the results based on only 8 subjects and the results obtained by the total experimental groups is remarkable and implicates a high degree of reliability of the obtained trends.

Another interesting aspect, aimed at illuminating how estimates of difficulty come about, can be reported in connection with the test "Number series" where experimental conditions additional to those described above were given. The items in "Number series" are not of multiple choice character like the items in "Levers" and "Synonyms", which was obviously the reason why certain subjects did not give any answers to certain items in the first-named test (this phenomenon never occurred in the other two tests). This happened altogether

25 times and every time subjects estimated the difficulty as "9" (very, very difficult). Whenever this was the case the experimenter explained to the subject how to solve the item in question and asked him anew to estimate the difficulty of the item.* As a new response "7" (difficult) was given 8 times, "8" (very difficult) 14 times, and only 4 times the answer remained unchanged "9" (very, very difficult). It would thus seem that those subjects who could not give an anser to an item responded purely in accordance with the experience of their own performance capacity ("if I can't solve this item it must be very, very difficult"). As soon as subjects have alternatives (as in the tests "Levers" and "Synonyms" through their multiple choice character) or are given one (as was the case in "Number series" when an item remained unsolved) there seems to be a tendency to estimate the difficulty of items which one actually cannot solve as somewhat lower than very, very difficult (there is a known chance of picking the right alternative, or the explanation is understood and causes a slight drop in estimated difficulty).

Other interesting features seen in Figures 3 and 4 are the slopes of the fitted regression lines. It is likely that knowledge about how and at what rate estimated difficulty increases with corresponding measurements based on performance could be useful for test constructing purposes together with other "subjective" and "objective" measurements.

In the further proceedings of the data analysis, the subjects participating in the three experiments were classified into subgroups homogeneous with respect to sex, age and performance on the test. The data of the various subgroups showed by and large the same general trends as have been described in the preceding sections; thus they need not be presented.

Another way to detect possible differences between estimates of subjects is to calculate, for each item in each test, the mean of estimates of subjects who actually solved an item correctly and for subjects who did not. Only items with at least 25% of observations in either of the categories "solved" or "unsolved" were considered. Figure 5 shows such means plotted against the order of items in the tests. Diagram A in Figure 5 represents data from "Number series", Diagram B those from "Levers", and Diagram C those from "Synonyms".

When examining the diagrams it is seen that subjects who solved the items of "Number series" and "Levers" correctly estimated the difficulty of the items throughout as lower than those subjects who did not, while no such tendency can be noticed in Diagram C, representing "Synonyms". The means of estimates of the two groups in each test were calculated. The differences between the central tendencies thus obtained (0.98 in "Number series", 0.78 in "Levers"

---

* These repeated estimates were not included when calculating the statistics presented in this report.

and 0.13 in "Synonyms") were statistically significant on the 1% level for "Number series" ($t = 2.982$, $df = 28$) and "Levers" ($t = 2.888$, $df = 26$) but not for "Synonyms". The reason why the estimates of the subjects with the "better" and the "poorer" performance do not differ with respect to the test "Synonyms" is probably found in the facts that this test consists of many "subjectively" easy items (the first 19 items out of 30 are estimated as less than medium difficult) and that also the performance on the test throughout is high.

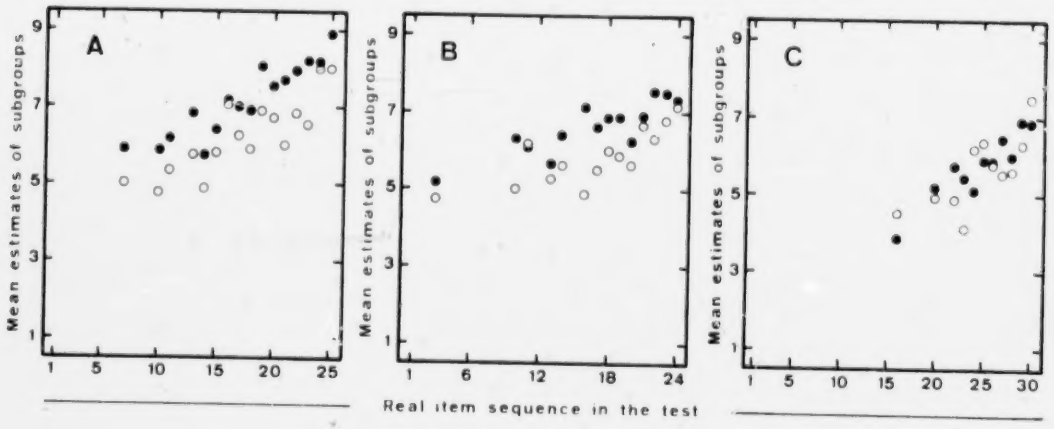


Fig. 5 Mean estimates of subjects who actually solved an item correctly (open circles) as well as mean estimates of subjects who did not (filled circles) plotted against the real order of occurrence in the tests. Diagram A shows data from "Number series", Diagram B those from "Levers" and Diagram C those from "Synonyms".

The main result of this study - the close and linear relationship between estimated difficulty of intellectual tasks and corresponding measurements based on performance - has now to be regarded as a rather well-established fact. The differences in the estimates of difficulty between the three tests seem mainly to be due to the different structures of objective item difficulty within each test. The number of given alternatives is probably the most important cue when estimating the difficulty of an item which one would not be able to answer when no alternatives are given. Given alternatives seem to cause a drop in estimated difficulty in such cases. It would be interesting to investigate to what degree a change in the number of alternatives in a given test causes differences in the estimates of difficulty and in the solution frequencies.

References

Borg, G. Interindividual scaling and perception of muscular force. Kungliga Fysiografiska Sällskapets i Lund Förhandlingar, 1961, 12, 117-125.

Borg, G. Physical performance and perceived exertion. Lund (Sweden): Gleerup, 1962.

Borg, G. Upplevd svårighet hos uppgifter i Ravens matriser. Unpublished manuscript, the Institute of Applied Psychology, the University of Stockholm, 1969 (in Swedish).

Borg, G. & Forsling, S. A psychological study on perceived difficulty. Educational and Psychological Research Bulletin, the University of Umeå, No. 1, 1964.

Borg, G. & Forsling, S. Bestämning av svårighetsgraden hos testuppgifter utifrån ett upplevelsekriterium. Rapport från pedagogisk-psykologiska institutionen, Umeå universitet, nr 3, 1965 (in Swedish).

Borg, G., Bratfisch, O. & Dornič, S. Perceived difficult of an immediate memory task. Reports from the Institute of Applied Psychology, the University of Stockholm, No. 15, 1971 (a).

Borg, G., Bratfisch, O. & Dornič, S. Perceived difficulty of a visual search task. Reports from the Institute of Applied Psychology, the University of Stockholm, No. 16, 1971 (b).

Borg, G., Bratfisch, O & Dornič, S. On the problems of perceived difficulty. Scandinavian Journal of Psychology, 1971, 12, 249-260 (c).

Bratfisch, O., Dornič. S. & Borg, G. Perceived difficulty of a motor-skill task as a function of training. Reports from the Institute of Applied Psychology, the University of Stockholm, No. 11, 1970.

Bratfisch, O., Dornič, S. & Borg, G. Perceived difficulty of items in a test of reasoning ability. Reports from the Institute of Applied Psychology, the University of Stockholm, No. 28, 1972.

Delta Battery Manual, Stockholm: Skandinaviska Testförlaget AB, 1970.

Ekman, G & Künnapas, T. The distribution function for similarity estimates. Reports from the Psychological Laboratories, the University of Stockholm, No. 286, 1969.

Guilford, J. P. & Cotzin, M. Judgement of difficulty of simple tasks. American Journal of Psychology, 1941, 54, 38-52.

Härnqvist, K. Individuella differenser och skoldifferentiering. SOU (Statens offentliga utredningar) 1960:13 (in Swedish).

Thurstone, L. L. & Thurstone, T. G. Factorial studies of intelligence. Chicago: University of Chicago Press, 1941 (Psychometric Monographs 2).

Abstract card

Bratfisch, O., Borg, G., and Dornič, S. Perceived item-difficulty in three tests of intellectual performance capacity. Reports from the Institute of Applied Psychology, the University of Stockholm, 1972, No. 29. - Three tests of intellectual performance capacity referring to factors V, S, and R according to Thurstone's system of primary mental abilities were administered to a total number of 34 subjects. Immediately after finishing an individual item subjects were asked to estimate the perceived difficulty of that item. The ratings were to be given on a symmetric scale with 9 categories with verbal expression labels. A high correlation between the rank order of items according to estimated difficulty and the real item sequence was obtained in all three tests used ($r \geq 0.92$). A linear relationship was found between estimated difficulty and standard scores corresponding to solution frequencies. A close correspondence was noticed between the widths and the levels of the ranges of the estimates on the one hand and the corresponding widths and levels of the standard scores ranges on the other hand. Subjects who could solve an item correctly tended to estimate the difficulty of that item as lower than subjects who could not.

Reference card

Bratfisch, O., Borg, G., and Dornič, S. Perceived item-difficulty in three tests of intellectual performance capacity. Reports from the Institute of Applied Psychology, the University of Stockholm, 1972, No. 29.

REPORTS SO FAR PUBLISHED IN THIS SERIES

No. 1, 1970 Borg, G.: RELATIVE RESPONSE AND STIMULUS SCALES

No. 2, 1970 Bratfisch, O.: TIME-ESTIMATIONS OF THE MAIN ACTIVITIES OF UNIVERSITY STUDENTS

No. 3, 1970 Hosman, J., and Borg, G.: THE MEAN AND STANDARD DEVIATION OF CROSS-MODALITY MATCHES: A STUDY OF INDIVIDUAL SCALING BEHAVIOR

No. 4, 1970 Borg, G., Edström, C-G., and Marklund, G.: A NEW METHOD TO DETERMINE THE EXPONENT FOR PERCEIVED FORCE IN PHYSICAL WORK

No. 5, 1970 Hosman, J.: THE FACTOR STRUCTURE OF MAGNITUDE PRODUCTIONS

No. 6, 1970 Hosman, J.: THE DIMENSIONALITY OF CROSS-MODALITY MATCHES

No. 7, 1970 Borg, G., and Hosman, J.: THE METRIC PROPERTIES OF ADVERBS

No. 8, 1970 Borg, G., Edgren, B., and Marklund, G.: A FLEXIBLE WORK TEST WITH A FEEDBACK SYSTEM GUIDING THE TEST COURSE

No. 9, 1970 Dornič, S., Künnapas, T., and Bratfisch, O.: SUBJECTIVE SIMILARITY AS A FUNCTION OF EXPOSURE TIME AND SHORT-TERM MEMORY

No. 10, 1970 Borg, G., Bratfisch, O., and Dornič, S.: ON PERCEIVED DIFFICULTY

No. 11, 1970 Bratfisch, O., Dornič, S., and Borg, G.: PERCEIVED DIFFICULTY OF A MOTOR-SKILL TASK AS A FUNCTION OF TRAINING

No. 12, 1970 Hosman, J., and Borg, G.: THE METRIC STRUCTURE OF VERBAL EXPRESSIONS: A FURTHER INVESTIGATION

No. 13, 1970 Valter, V.: DEDUCTION AND VERIFICATION OF A QUANTUM PSYCHOPHYSICAL EQUATION

No. 14, 1970 Dornič, S., Bratfisch, O., and Santesson, A.: VERBAL FACTOR IN IMMEDIATE MEMORY

No. 15, 1971 Borg, G., Bratfisch, O., and Dornič, S.: PERCEIVED DIFFICULTY OF AN IMMEDIATE MEMORY TASK

No. 16, 1971 Borg, G., Bratfisch, O., Dornič, S.: PERCEIVED DIFFICULTY OF A VISUAL SEARCH TASK

No. 17, 1971 Borg, G.: A NOTE ON A "DISPERSION METHOD" IN PSYCHOPHYSICAL SCALING